

The Diagnostic and Prognostic Value of Neuropsychological Assessment in Memory Clinic Patients

Citation for published version (APA):

Jansen, W. J., Handels, R. L. H., Visser, P. J., Aalten, P., Bouwman, F., Claassen, J., van Domburg, P., Hoff, E., Hoogmoed, J., Leentjens, A. F. G., Rikkert, M. O., Oleksik, A. M., Smid, M., Scheltens, P., Wolfs, C., Verhey, F., & Ramakers, I. H. G. B. (2017). The Diagnostic and Prognostic Value of Neuropsychological Assessment in Memory Clinic Patients. *Journal of Alzheimer's Disease*, 55(2), 679-689. <https://doi.org/10.3233/JAD-160126>

Document status and date:

Published: 01/01/2017

DOI:

[10.3233/JAD-160126](https://doi.org/10.3233/JAD-160126)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

The Diagnostic and Prognostic Value of Neuropsychological Assessment in Memory Clinic Patients

Willemijn J. Jansen^{a,*}, Ron L.H. Handels^a, Pieter Jelle Visser^{a,b}, Pauline Aalten^a, Femke Bouwman^b, Jurgen Claassen^c, Peter van Domburg^d, Erik Hoff^e, Jan Hoogmoed^{f,g}, Albert F.G. Leentjens^a, Marcel Olde Rikkert^c, Ania M. Oleksik^h, Machiel Smidⁱ, Philip Scheltens^b, Claire Wolfs^a, Frans Verhey^a and Inez H.G.B. Ramakers^a

^aDepartment of Psychiatry and Neuropsychology, Alzheimer Center Limburg, School for Mental Health and Neuroscience, Maastricht University, Maastricht, The Netherlands

^bDepartment of Neurology, VUmc Alzheimer Center, VU University Medical Center, Neuroscience Campus Amsterdam, Amsterdam, The Netherlands

^cDepartment of Geriatrics, Radboud Alzheimer Center, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands

^dDepartment of Neurology, Zuyderland Medical Center, Sittard, The Netherlands

^eDepartment of Neurology, Zuyderland Medical Center, Heerlen, The Netherlands

^fDepartment of Geriatrics, Laurentius Hospital, Roermond, The Netherlands

^gDepartment of Geriatrics, St. Jans Gasthuis, Weert, The Netherlands

^hDepartment of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands

ⁱDepartment of Geriatrics, Zuyderland Medical Center, Sittard, The Netherlands

Accepted 29 August 2016

Abstract.

Background: Neuropsychological testing has long been embedded in daily clinical practice at memory clinics but the added value of a complete neuropsychological assessment (NPA) to standard clinical evaluation is unknown.

Objective: To evaluate the added diagnostic and prognostic value of NPA to clinical evaluation only in memory clinic patients.

Methods: In 221 memory clinic patients of a prospective cohort study, clinical experts diagnosed clinical syndrome (subjective cognitive impairment (SCI), mild cognitive impairment (MCI), or dementia) and etiology (Alzheimer's disease (AD) or no AD), and provided a prognosis of disease course (decline or no decline) before and after results of NPA were made available. The reference standard was a panel consensus based on all clinical information at baseline and up to 2 follow-up assessments.

Results: With NPA data available, clinicians changed their initial syndromal diagnosis in 22% of patients, and the etiological diagnosis as well as the prognosis in 15%. This led to an increase in correctly classified cases of 18% for syndromal diagnosis, 5% for etiological diagnosis, and 1% for prognosis. NPA data resulted in the largest improvement in patients initially classified as SCI (syndrome: 93.3% ($n = 14$) correctly reclassified, etiology: net reclassification improvement [NRI] = 0.61, prognosis: NRI = 0.13) or MCI (syndrome: 89.3% ($n = 23$) correctly reclassified, etiology: NRI = 0.17, prognosis: NRI = 0.14), while there was no improvement in patients with dementia (syndrome: 100% ($n = 1$) correctly reclassified, etiology: NRI = -0.05, prognosis: NRI = -0.06). Overall, inclusion of NPA in the diagnostic process increased confidence in all diagnoses with 6–7%.

*Correspondence to: Willemijn J. Jansen, Department of Psychiatry and Neuropsychology, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Tel.: +31 43 3884089; Fax: +31 43 3884092; E-mail: willemijn.jansen@maastrichtuniversity.nl.

Conclusion: Administration of a complete NPA after standard clinical evaluation has added value for diagnosing cognitive syndrome and its underlying etiology in patients regarded as non-demented based on the first clinical impression.

Keywords: Alzheimer's disease, cognitive disorders, consensus, diagnosis, prognosis, mild cognitive impairment, neuropsychological tests, outpatient clinic, reclassification

INTRODUCTION

A timely and accurate diagnosis in people with cognitive complaints can answer concerns about symptoms [1], facilitate adjustment [2], and is essential to initiate appropriate care and treatment [3]. However, differentiating between normal aging, mild cognitive impairment (MCI), and dementia, determining the underlying cause for cognitive complaints, and making a reliable prognosis for a patient is often challenging in clinical practice [4, 5].

The diagnostic assessment of patients presenting at a memory clinic with a cognitive complaint starts with clinical evaluation by a medical doctor and may be followed by neuropsychological assessment (NPA), depending on the setting. Many studies have shown that several single neuropsychological tests, isolated or combined, can predict future dementia [6] and differentiate between Alzheimer's disease (AD) and other common causes for cognitive impairment [7–10]. Neuropsychological testing has long been embedded for both diagnostic and treatment purposes in daily clinical practice at memory clinics and is a useful additional test according to clinical criteria for AD and neurocognitive disorders [11]. However, the added diagnostic and prognostic value of a complete NPA to the clinical evaluation has not been systematically assessed. Identification of the optimal diagnostic strategy in patients with cognitive complaints is important to ensure an accurate diagnosis and prognosis, without preventing valuable testing or exposing patients to unnecessary testing.

We therefore aimed to evaluate the added diagnostic and prognostic value of NPA as administered after the clinical evaluation in tertiary memory clinic patients and to derive strategies for optimal use of NPA as a diagnostic instrument. A staged clinical decision-making process was used to compare the usefulness of the standard clinical evaluation without and with NPA to diagnose cognitive syndrome, differentiate between AD and other etiologies, and to make a disease course prognosis. We examined change in diagnostic accuracy, correct reclassification, net reclassification improvement and change in diagnostic confidence related to NPA when added to the standard clinical evaluation.

MATERIALS AND METHODS

Participants

We selected participants from the prospective Leiden-Alzheimer Research Netherlands (LeARN) study [12]. In this study, consecutive patients from four Dutch academic memory clinics (Maastricht University Medical Center, VU Medical Center, Radboud University Medical Center, Leiden University Medical Center) were included between 2009 and 2011 and annually followed for two years. Inclusion criteria were referral for a cognitive complaint by a physician or suspicion of a cognitive disorder, a Mini-Mental State Examination (MMSE [13]) score ≥ 20 [14], and a Clinical Dementia Rating (CDR [15]) score ≤ 1 . Patients were excluded if the cognitive problems were due to stroke, neurological disorder, major psychiatric disorder, or substance abuse, or when a reliable informant was absent. The Medical Ethics Committee of each participating center approved this study and all patients initially included and their informal caregivers gave written informed consent.

Of the 304 participants who met inclusion criteria for the LeARN study, 221 patients were included in the current analyses. Eligibility for these analyses required availability of at least one follow-up assessment at one or two years after the baseline assessment. Patients who died ($n = 3$), were institutionalized ($n = 4$), unwilling to further participate ($n = 62$), or could not be reached ($n = 14$) were therefore excluded. Of these, 163 (74%) had both a 1-year and a 2-year follow-up assessment, 39 (18%) had only a 1-year and 19 (9%) had only a 2-year follow-up assessment. Eligible patients and excluded patients did not differ in age at baseline, years of education, MMSE score, CDR score, or clinical diagnosis at baseline.

Clinical evaluation

At baseline and follow-up, participants underwent a standardized clinical evaluation consisting of a detailed history provided by the patient and informant, a psychiatric, neurological, and physical

examination, and assessment using clinical rating scales. Clinical rating scales included the MMSE [13], the CDR [15], the Geriatric Depression Scale-15 (GDS-15) [16], the Neuropsychiatric Inventory (NPI) [17], and the Disability Assessment for Dementia (DAD) [18].

Neuropsychological assessment

Neuropsychological assessment was administered at baseline and repeated annually, and consisted of a standardized battery of Dutch versions of cognitive tests [19]. Tests included the Auditory Verbal Learning Test (AVLT) [20] comprising an immediate recall of 5 trials of 15 semantically unrelated words and a delayed recall and recognition after 20 minutes to assess verbal memory, the Visual Association Test (VAT) [21] to assess visual memory, the Digit Span task of the WAIS III (forward and backward) [22] to evaluate working memory, a one-minute verbal fluency test using a semantic category cue (animals) [23] to assess language function, the Letter Digit Substitution Test (LDST) [24] to assess information processing speed, the Stroop Color-Word Test (SCWT) [25] and the Trail Making Test (TMT) (Part A 'letters' and Part B 'concept shifting') [26] to assess attention and executive functioning, and subtests of the Visual and Object Spatial Perception (VOSP) test [27] to assess visual perception. Raw test scores were converted to z-scores using Dutch norms adjusting for age, gender, and education.

Neuroimaging

Neuroimaging by magnetic resonance imaging (MRI) took place at baseline only. Measures of medial temporal lobe atrophy (MTA), scored with the MTA scale [28], white matter lesions scored with the Fazekas scale [29], and the number of infarcts, lacunes and microbleeds were derived from the 3T MRI images, which were acquired following the standardized Parelsnoer protocol [19].

Diagnostic procedure

All patient information was summarized in digital case descriptions and presented to individual members of an expert panel in a stepwise fashion to simulate clinical practice (Fig. 1), as previously described [30]. An expert panel was composed of

clinicians from three different disciplines (geriatrics, geriatric psychiatry, and neurology) with at least five years of clinical experience in a memory clinic. This included experience with integrating neuropsychological test information in their diagnostic reasoning. At each stage, experts were asked to individually diagnose the patients' syndrome and etiology underlying the cognitive complaints, and to predict the course of cognitive symptoms and daily functioning in the next two years. The syndrome could be classified as: subjective cognitive impairment (SCI), MCI, or dementia; the etiology as: AD, vascular, frontotemporal, Lewy Body or Parkinson's, other, or no neurodegenerative disease; and the disease course prognosis as: improvement, stable, or decline. Experts also indicated confidence in their diagnosis on a scale from 0% (very uncertain) to 100% (fully convinced) at each stage. Firstly, the information from the clinical evaluation as described above was presented to the individual experts resulting in the initial *clinic only* diagnosis on syndrome, etiology and predicted disease course. Secondly, neuropsychological test results (z-scores) and a short summary of the neuropsychologist were presented as add-on to the clinical information and interpreted by the experts to make a subsequent *clinic + NPA* diagnosis. The short summary provided by the neuropsychologist consisted of 2-3 sentences in which performance on all individual cognitive domains tested was objectively described as "(very) high", "above average", "average", "below average", or "(very) low" without including an opinion or diagnosis. Characteristic observations conducted during testing were also included. Thirdly, the baseline information on the clinical evaluation and NPA results combined with measures from the MRI scan was presented and integrated by the experts in the baseline diagnosis. Lastly, all information from baseline and follow-up was presented to determine the follow-up diagnosis.

The same trio of experts completed all steps for each case. Experts did not take notes during any of the assessments. Experts were not obliged to apply strict decision rules but were recommended to use diagnostic guidelines as they would do in clinical practice. The patient's name was de-identified and the number of cases in which the expert and the participant originated from the same clinical center was minimized to prevent experts from recognizing a case. However, this could not be fully prevented, which resulted in the recognition of one case by one expert.

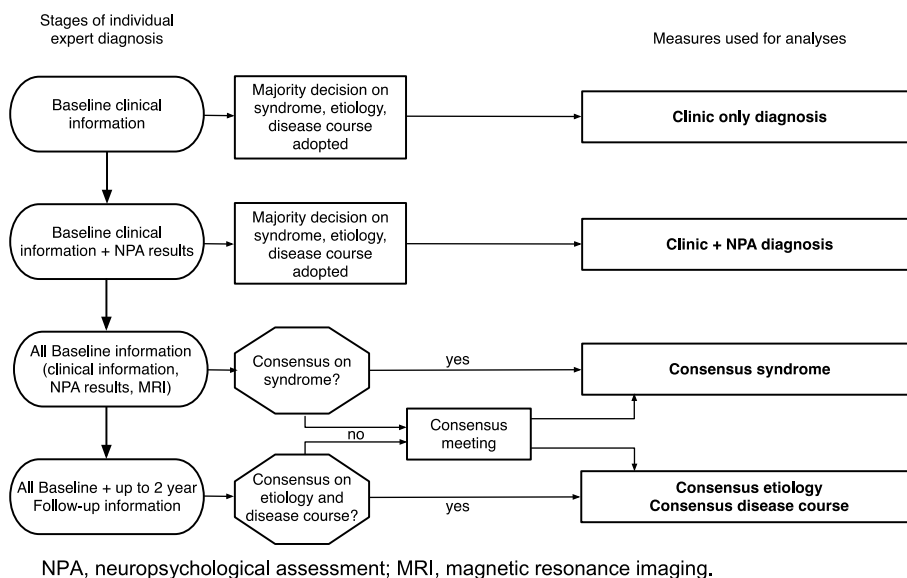


Fig. 1. The staged diagnostic procedure followed for each patient. NPA, neuropsychological assessment; MRI, magnetic resonance imaging.

Consensus procedure

For the *clinic only* and *clinic + NPA* diagnoses for each patient, we adopted the syndrome (SCI, MCI, or dementia), etiology (dichotomized into AD or no AD) and prognosis (dichotomized into decline or no decline) as designated by the majority of experts (at least 2 out of 3).

For the baseline diagnoses, consensus among members of an expert panel on the syndrome was determined and for the follow-up diagnoses, consensus on the etiology and predicted disease course was determined. In case of discrepancy, a meeting was organized approximately 2 weeks after the initial rating in which the expert trios were invited to express their arguments and reach consensus. This procedure resulted in the *consensus syndrome*, *consensus etiology* (dichotomized), and *consensus disease course* (dichotomized) that were used as reference standards for syndrome, etiology and disease course prognosis.

Data analysis

Syndrome, etiology and prognosis for the *clinic only* and *clinic + NPA* diagnoses were compared against their reference standards. Percentages correctly classified in the *clinic only* and *clinic + NPA* diagnoses were calculated for syndrome, etiology and predicted disease course and compared with the McNemar test statistic for paired proportions. We calculated sensitivity and specificity, positive predictive

value and negative predictive value for the etiology and disease course classifications of the *clinic only* and *clinic + NPA* diagnoses. Mean certainty of the *clinic only* and *clinic + NPA* diagnoses and prognoses was compared with paired-sample *t*-tests. The added value of NPA was expressed in percentages of (correct) reclassifications in syndrome, etiology and predicted disease course from *clinic only* and *clinic + NPA* diagnosis determined using reclassification tables. Category-based Net Reclassification Indices (NRI) for events (AD or decline) and non-events (no AD or no decline) were calculated for the etiology and predicted disease course classifications. The NRI is the difference in proportions reclassified upwards and downwards among events versus non-events, or $NRI = [Pr(up | events) - Pr(down | events)] + [Pr(down | non-events) - Pr(up | non-events)] = NRI \text{ events} + NRI \text{ non-events}$ [31]. NRI events and NRI non-events represent the net percentage of patients with or without events correctly assigned after receiving the NPA results. Z-scores for the NRIs were calculated according to the method of Pencina et al. [31] and compared with the standard normal distribution. To identify in which cases the NPA had added value, clinical characteristics and NPA results of patients in whom syndrome, etiology, or course was correctly reclassified from *clinic only* to *clinic + NPA* diagnosis were compared with those of patients in whom there was no or incorrect reclassification using independent *t*-tests and chi-square tests. Percentages of correct and incorrect

reclassification were calculated for each value of the predictor. We also characterized patients in whom etiology or course was correctly not reclassified from *clinic only* to *clinic + NPA* diagnosis to identify in which cases the NPA had no added value. Analyses were conducted with SPSS version 20.0 (Chicago, IL, USA) and significance level was set at $p < 0.05$.

RESULTS

Baseline characteristics of the sample are displayed in Table 1. Of the 221 patients included in the analyses, 52 were diagnosed with SCI, 90 with MCI, and 79 with dementia at *consensus syndrome*. Based on the *consensus etiology*, cognitive problems were caused by AD in 102 (46%) patients (2 SCI, 43 MCI, 57 dementia). Clinically relevant cognitive or functional decline over the follow-up period was observed in 128 (58%) participants (5 SCI, 53 MCI, 70 dementia) based on the *consensus disease course*. The experts did not reach consensus on the reference etiology in 1 case, and on the reference disease course in 3 cases. In these participants a majority decision was adopted.

Comparison of *clinic only* and *clinic + NPA* diagnoses

In these analyses, we directly compared the syndrome, etiology, and predicted disease course assigned to patients in the *clinic only* and in the *clinic + NPA* diagnoses with the reference standards to examine correct classification. The syndrome was initially correctly classified in 70.7% ($n = 152$, indeterminate diagnosis in 6 cases) of patients, which

increased to 88.5% ($n = 192$, indeterminate diagnosis in 4 cases) after NPA results were disclosed ($\chi^2 = 15.2$, $p < 0.001$). The mean certainty of the experts in their syndrome classification increased from 68.0% to 74.3% ($t = -10.9$, $p < 0.001$) after making NPA data available. Diagnostic accuracy measures of the etiology and predicted disease course classifications are displayed in Table 2. Correct classifications of etiology in 76.5% of patients by the *clinic only* diagnosis increased to 81.0% by the *clinic + NPA* diagnosis ($\chi^2 = 10.0$, $p = 0.002$). In making a disease course prognosis, the percentage correctly classified cases slightly increased after presentation of NPA results from 74.7% to 75.6% ($\chi^2 = 0.5$, $p = 0.48$). The experts were 6–7% more certain of both etiology ($t = -12.1$, $p < 0.001$) and predicted disease course ($t = -10.1$, $p < 0.001$) classifications after reviewing the NPA. Supplementary Table 2 presents measures of classification separately for patients with up to 2 years of follow-up and patients with only a 1-year follow-up assessment.

Reclassification of diagnoses

We also examined the reclassification of syndrome, etiology, and predicted disease course from *clinic only* diagnosis to *clinic + NPA* diagnosis and determined whether this reclassification was correct based on the reference standards. Table 3 displays the reclassification of syndromes assigned to patients at the *clinic only* and *clinic + NPA* stages compared to the *consensus syndrome*. Overall, the initial classification of syndrome changed in 20.7% ($n = 44$, indeterminate diagnosis in 9 cases) after reviewing NPA results, which were correct when compared to

Table 1
Baseline characteristics

Characteristics	Consensus syndrome			
	All ($n = 221$)	SCI ($n = 52$)	MCI ($n = 90$)	Dementia ($n = 79$)
Age, years	66.6 \pm 9.3	62.9 \pm 9.0	67.4 \pm 8.7	68.0 \pm 9.7
Female, n (%)	78 (35.3)	20 (38.5)	30 (33.3)	28 (35.4)
Education, years	11.2 \pm 3.4	11.8 \pm 3.6	11.2 \pm 3.5	10.7 \pm 3.3
MMSE score	25.9 \pm 3.1	28.4 \pm 1.7	26.7 \pm 2.5	23.5 \pm 2.7
CDR score	0.6 \pm 0.3	0.3 \pm 0.2	0.6 \pm 0.3	0.8 \pm 0.3
GDS-15 score	3.3 \pm 2.8	3.4 \pm 2.5	3.2 \pm 2.6	3.3 \pm 3.1
NPI score	14.5 \pm 14.0	12.5 \pm 12.1	12.7 \pm 14.0	17.8 \pm 14.7
DAD score	86.2 \pm 15.9	93.4 \pm 9.6	89.9 \pm 13.1	76.7 \pm 17.9
AD [§] , no. (%)	102 (46.2)	2 (3.8)	43 (47.8)	57 (72.2)
Declined [§] , n (%)	128 (57.9)	5 (9.6)	53 (58.9)	70 (88.6)

[§]according to the consensus diagnosis. Data are mean \pm SD or number (%). SCI, subjective cognitive impairment; MCI, mild cognitive impairment; MMSE, Mini-Mental State Examination; CDR, Clinical Dementia Rating scale; GDS, Geriatric Depression Scale; NPI, Neuropsychiatric Inventory; DAD, Disability Assessment for Dementia; AD, Alzheimer's disease.

Table 2

Accuracy measures of *clinic only* and *clinic + NPA* classifications of etiology and disease course compared to *consensus etiology* and *consensus disease course*

Reference consensus	Stage	TP	FP	TN	FN	Sensitivity	Specificity	PPV	NPV	Correctly classified	Mean certainty
Etiology (AD y/n)	Clinic only	71	21	98	31	69.6%	82.4%	77.2%	76.0%	76.5%	57.7%
	Clinic + NPA	80	20	99	22	78.4%	83.2%	80.0%	81.8%	81.0%*	64.5%**
Disease course (decline y/n)	Clinic only	109	37	56	19	85.2%	60.2%	74.7%	74.7%	74.7%	62.8%
	Clinic + NPA	114	40	53	14	89.1%	57.0%	74.0%	79.1%	75.6%	68.5%**

* $p < 0.01$. ** $p < 0.001$. TP, true positives; FP, false positives; TN, true negatives; FN, false negatives; PPV, positive predictive value; NPV, negative predictive value; AD, Alzheimer's disease; NPA, neuropsychological assessment.

Table 3

Reclassification of syndrome at *clinic only* and *clinic + NPA* stages compared to the *consensus syndrome* for patients classified as SCI, MCI, or dementia

Stage			
Clinic only*	Clinic + NPA*	Consensus syndrome	Number of patients (%#)
SCI ($n = 45$)	SCI ($n = 30$)	SCI	28 (62.2)
		MCI	2 (4.4)
		Dementia	0 (0.0)
	MCI ($n = 15$)	SCI	1 (2.2)
		MCI	14 (31.1)
		Dementia	0 (0.0)
MCI ($n = 109$)	SCI ($n = 18$)	SCI	16 (14.7)
		MCI	2 (1.8)
		Dementia	0 (0.0)
	MCI ($n = 78$)	SCI	3 (5.8)
		MCI	62 (56.8)
		Dementia	13 (11.9)
Dementia ($n = 61$)	Dementia ($n = 10$)	SCI	0 (0.0)
		MCI	1 (0.9)
		Dementia	9 (8.3)
	Indeterminate ($n = 3$)	SCI	2 (1.8)
		MCI	1 (0.9)
		Dementia	0 (0.0)
Indeterminate ($n = 6$)	MCI ($n = 1$)	SCI	0 (0.0)
		MCI	1 (1.6)
		Dementia	0 (0.0)
	Dementia ($n = 60$)	SCI	0 (0.0)
		MCI	3 (9.8)
		Dementia	57 (93.4)
Indeterminate ($n = 6$)	SCI ($n = 2$)	SCI	2 (33.3)
		MCI	0 (0.0)
		Dementia	0 (0.0)
	MCI ($n = 3$)	SCI	0 (0.0)
		MCI	3 (50.0)
		Dementia	0 (0.0)
Indeterminate ($n = 1$)	Indeterminate ($n = 1$)	SCI	0 (0.0)
		MCI	1 (16.7)
		Dementia	0 (0.0)

*Majority rating. Supplementary Figure 2 shows the distribution of majority (2/3 or 3/3) and indeterminate syndrome diagnosis ratings. # % of Clinic only. NPA, neuropsychological assessment; SCI, subjective cognitive impairment; MCI, mild cognitive impairment.

the *consensus syndrome* in 90.9% ($n = 40$). In cases initially assigned as SCI, 33.3% ($n = 15$) were reclassified to MCI of which correctly in 93.3% ($n = 14$).

Of the patients who were initially diagnosed with the MCI syndrome, 25.7% ($n = 28$) were reclassified to SCI or dementia, for whom the reclassification was

Table 4
Measures of reclassification in etiology and disease course from *clinic only* to *clinic + NPA* diagnosis

Reference consensus	Syndrome at clinic only diagnosis	Reclassified after NPA	Of which correctly	NRI events	NRI non-events	NRI
Etiology (AD y/n)	All (<i>n</i> = 221) [#]	32 (14.5%)	21 (65.6%)	8.8%*	0.8%	0.10*
	SCI (<i>n</i> = 45)	10 (22.2%)	6 (60.0%)	66.7%*	-5.1%	0.61*
	MCI (<i>n</i> = 109)	13 (11.9%)	11 (84.6%)	9.6%*	7.0%	0.17*
	Dementia (<i>n</i> = 61)	7 (11.4%)	3 (42.9%)	0.0%	-5.3%	-0.05
Disease course (decline y/n)	All (<i>n</i> = 221) [#]	34 (15.4%)	18 (52.9%)	3.9%	-3.2%	0.01
	SCI (<i>n</i> = 45)	10 (22.2%)	3 (30.0%)	28.6%	-15.8%*	0.13
	MCI (<i>n</i> = 109)	18 (16.5%)	13 (72.2%)	7.9%*	6.5%	0.14*
	Dementia (<i>n</i> = 61)	3 (4.9%)	0 (0.0%)	-5.6%*	0.0%	-0.06*

* $p < 0.05$. [#]Total number includes 6 indeterminate cases in which no majority syndrome could be assigned. NPA, neuropsychological assessment; NRI, net reclassification improvement; AD, Alzheimer's disease; SCI, subjective cognitive impairment; MCI, mild cognitive impairment.

correct in 89.3% (*n* = 25). Only 1 case with initial dementia was reclassified to MCI, which matched the *consensus syndrome*.

Etiology was reclassified by the clinical experts in the *clinic + NPA* diagnosis in 14.5% (*n* = 32), which were correct when compared to the *consensus etiology* in 65.6% (*n* = 21, Table 4). Overall, there was net improvement in diagnosing etiology (NRI = 0.10, $z = 1.85$, $p = 0.032$) after NPA and this was mainly caused by an improvement of 8.8% in reclassifying the etiology of AD cases from 'no AD' to 'AD'.

The disease course prognosis was reclassified in 15.4% (*n* = 34) of patients, of which correctly according to the *consensus prognosis* in 52.9% (*n* = 18). The overall NRI for predicted disease course was 0.01 ($z = 0.12$, $p > 0.05$) indicating no significant improvement.

To evaluate the added value of NPA across different syndromes, we examined reclassification from *clinic only* to *clinic + NPA* diagnosis according to the *clinic only* diagnosis of syndrome. Etiology was more often correctly than incorrectly reclassified in patients with an initial SCI or MCI syndrome, but not dementia. NRIs were different per diagnostic group and significant for the etiology diagnosis in patients with initial SCI (NRI = 0.61, $z = 1.81$, $p = 0.036$) and MCI (NRI = 0.17, $z = 2.50$, $p = 0.006$), and for the disease course prognosis in patients with initial MCI (NRI = 0.14, $z = 1.68$, $p = 0.047$). These improvements were all driven by the NRI events, indicating that the NPA is beneficial for ruling in AD etiology. For the disease course prognosis, there were more correct than incorrect reclassifications in patients with initial MCI, but not SCI and dementia. In initial dementia patients, a net worsening of classification (NRI = -0.06, $z = -1.73$, $p = 0.042$) due to the NPA was found in making a disease course prognosis.

Indicators of correct reclassification

Characteristics of participants in whom syndrome, etiology, or predicted disease course were correctly reclassified or correctly not reclassified (i.e., unchanged correct classification) after NPA are displayed in Supplementary Table 1. Higher MMSE (27.5 versus 25.6, $t = -3.5$, $p = 0.001$) and lower CDR scores (0.5 versus 0.6, $t = 3.4$, $p = 0.001$) were found to be indicative of correct syndrome reclassification. Correct reclassification of etiology was characterized by worse AVLT delayed recall *z*-score (-2.2 versus -1.5, $t = 2.1$, $p = 0.039$) and better performance on digit span (15.7 versus 13.1, $t = -2.5$, $p = 0.012$). Further, lower CDR scores (0.4 versus 0.6, $t = -2.6$, $p = 0.010$) and higher AVLT immediate recall *z*-scores (-0.7 versus -1.5, $t = -2.6$, $p = 0.011$) were indicators of correct reclassification of predicted disease course. Figure 2 depicts the distribution of correct and incorrect reclassifications for values of the neuropsychological tests and scales that were related to correct reclassification and Supplementary Figure 1 further specifies the direction of these reclassifications. Patients in whom the NPA did not change an initially correct classification were in general characterized by worse performance on both clinical rating scales and neuropsychological tests. Patients' age, years of education, score on a depression or disability questionnaire, and other neuropsychological test results were not indicative of correct reclassification or correct no reclassification.

DISCUSSION

In this study, we assessed the added value of NPA to standard clinical evaluation in the diagnostic process in memory clinic patients. NPA changed the diagno-

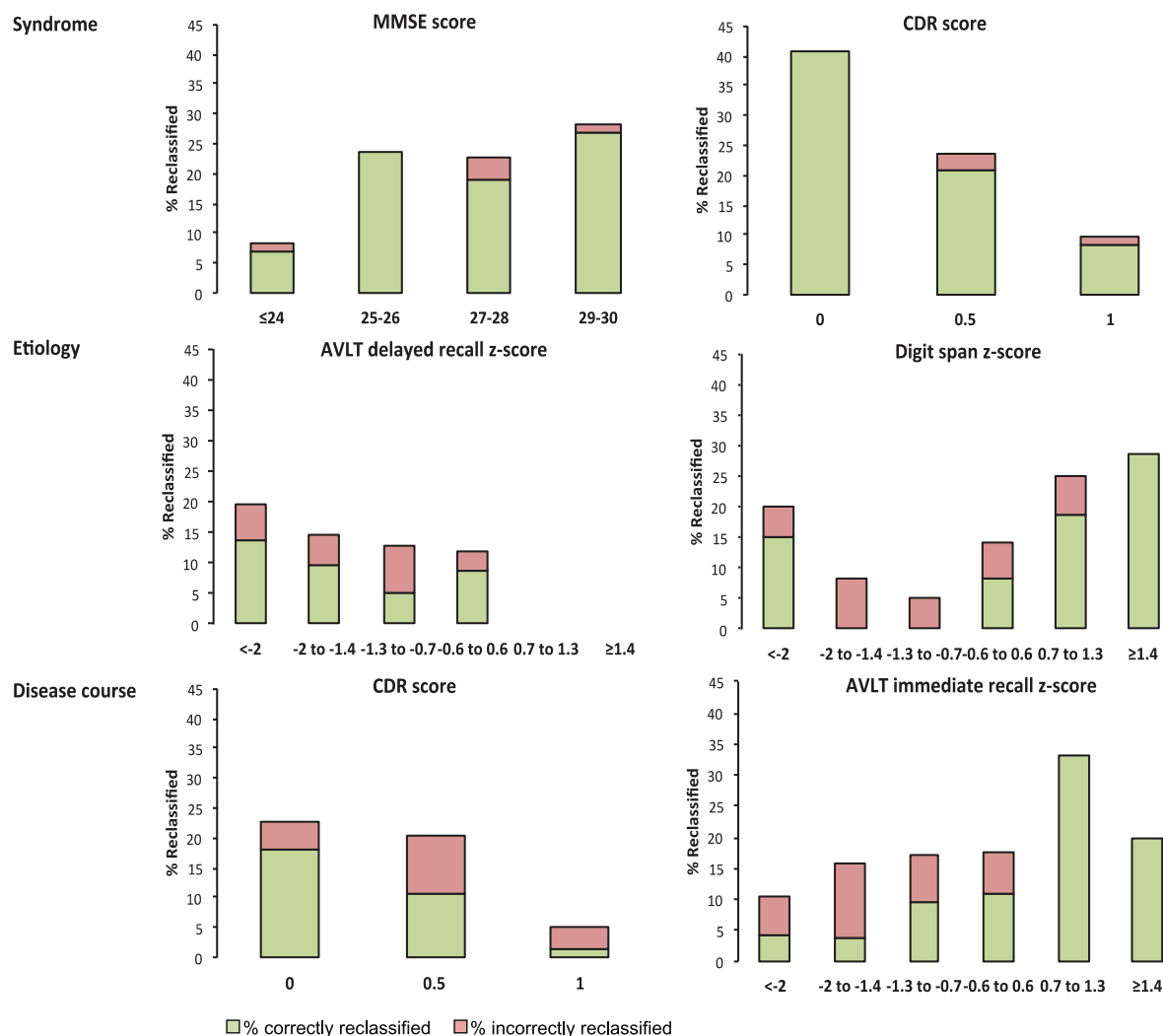


Fig. 2. Percentages correctly and incorrectly reclassified from *clinic only* to *clinic + NPA* diagnosis. MMSE, Mini-Mental State Examination; AVLT, Auditory Verbal Learning Task; CDR, Clinical Dementia Rating scale.

sis of the cognitive syndrome in 22% of patients and the underlying etiology as well as the disease course prognosis in 15%. Overall, this led to an increase in correctly classified cases that could be attributed to the NPA of about 18% for the cognitive syndrome, 5% for the underlying etiology, and 1% for the predicted disease course. Results demonstrated that the diagnostic and prognostic value of NPA when added to the clinical evaluation depended on which initial syndrome was diagnosed. The NPA shows added value for diagnosing syndrome and etiology in patients with initial SCI or MCI, and for predicting disease course in patients with an initial MCI syndrome. There was no added diagnostic or prognostic value of the NPA in patients suspected of dementia after the initial clinical evaluation. Further, inclusion of the NPA in the

diagnostic process increased confidence in the diagnosis.

NPA had high added value in diagnosing the cognitive syndrome, which is in line with the general idea that neuropsychological tests are accurate in determining the severity of cognitive deficits [32]. The value of NPA in diagnosing syndrome did not depend on initial under- or overestimation of cognitive complaints. There was added value of NPA in diagnosing etiology in patients with an initial SCI or MCI syndrome, and for predicting disease course in patients with initial MCI. Our overall NRI was mainly directed by the events NRI in these patients, indicating that the NPA supported the detection of persons who had AD or declined at follow-up. In patients initially suspected of a dementia syndrome, the added

value of the NPA was limited, and in some cases the NPA even misled the experts into believing patients would remain stable. Furthermore, incorrect reclassifications of predicted disease course often occurred in patients with SCI in whom it is likely to take more than two years before a disease is clinically expressed [33]. When determining whether or not to conduct a NPA or to reclassify after the NPA has been conducted, tests displayed in Fig. 2 and Supplementary Figure 1 can be used as guideline in decision-making. Results also showed that age and education, depression and disability scores were unrelated to the impact of reclassification after NPA on the diagnosis. This indicates that the added value of NPA in the diagnostic process of memory clinic patients is not dependent on these factors.

To our knowledge, no earlier study investigated the added value of a complete NPA to the standard clinical evaluation or examined neuropsychological test results in a memory clinic setting by verifying diagnoses through a reference diagnosis such that correct reclassifications could be determined. Other studies concerning dementia diagnostics found changes caused by the NPA of 11% for the etiology diagnosis [34], 25% for diagnosing a combination of syndrome and etiology [35], and 26% for diagnosing a combination of syndrome and etiology when both NPA and MRI were added to the standard clinical assessment [36]. However, divergent study populations and classification categories prevented reliable comparison with our study.

This study simulated clinical decision-making using comprehensive case descriptions and panel discussions. This combination of individual and plenary approaches has been found to produce very similar results to a full plenary approach as often adopted in clinical practice [37]. Additionally, assigning diagnoses after each step allowed us to compare classifications but also to investigate reclassifications, which are important to consider when the goal is to gain insight in the added value of a diagnostic instrument. The NRI measure is implicitly weighted by the event rate [38], which is useful when certain outcomes are not as prevalent as others, as is the case in this study. Another strength of this study is that diagnoses were compared against a reference standard to assess correctness. It should be noted that clinical applications of the NPA other than diagnostics, such as psycho-education, cognitive training and treatment, were not assessed here while these are also important utilities of NPA [39].

A limitation of the study design is that incorporation bias may have occurred since the index test results were also part of the reference standard, as is often inevitable in dementia diagnostic studies [40]. This may have led to an overestimation of the various measures of diagnostic accuracy. Also, there was a maximum follow-up of only two years. There might have been an overestimation of expected decline in dementia patients while the period is too short to decline for persons with SCI. Most patients were followed for two years but some patients only received a one-year follow-up assessment, which slightly affected the diagnostic accuracy measures. Further, although amyloid and tau biomarkers are more often used in clinical practice, we did not incorporate this information in the diagnostic process nor did we verify the etiology reference standard based on biomarker information. Moreover, the expert panels already diagnosed etiology and predicted disease course correctly in about 75% based on the clinical evaluation only. This high rate is likely explained by the extensive information made available in the clinical evaluation leaving less room for improvement for NPA and by the extensive diagnostic experience of our clinical experts. Additionally, this study was performed in specialized centers and results may therefore not be generalizable to other settings. Use of multidisciplinary teams including a neuropsychologist who interprets and transfers NPA findings in a comprehensive manner during the diagnostic process may overcome this limitation in less specialized centers. The added value of NPA may also depend on the choice of NPA tests or the order of presentation of additional test results, and may prove differently when more sensitive or specific measures would be used or when NPA and MRI would be presented in a different order. Lastly, our simulation was a simplification of actual practice because we presented summarized NPA results, which possibly led to an underestimation of the impact of the NPA since a neuropsychologist would also provide a contextual interpretation of the results, and take part in multidisciplinary meetings in clinical practice.

Although the NPA is widely applied in patients presenting with cognitive complaints, the diagnostic and prognostic value was not evaluated before in a large prospective study simulating clinical practice. We conclude that the added value of the NPA depends on the initial clinical impression of syndrome, and is different for diagnosing syndrome, etiology or predicting disease course. This may stimulate a more individualized approach in the diagnostic evaluation

of persons with cognitive complaints, which may benefit cost-effectiveness of NPA. To make an optimal diagnosis and prognosis, we recommend the use of NPA as decisive investigation in patients who are considered non-demented at initial clinical evaluation and in patients in whom more certainty about the diagnosis is desired. In patients suspected of dementia, the diagnostic value of the NPA is limited and this should lead to reconsideration of administering NPA bearing in mind alternative purposes of NPA such as inventorying profiles of cognitive strengths and weaknesses, burden and care needs.

ACKNOWLEDGMENTS

This research is performed within the framework of CTMM, the Center for Translational Molecular Medicine (<http://www.ctmm.nl>), project LeARN (grant 02 N-101). The funding body had no role in the study design, the collection, analysis or interpretation of data, or the writing of the manuscript. The funding body checked the manuscript for possible infringement of intellectual property rights and approved the manuscript for publication without suggesting any revisions.

The authors thank Nico Rozendaal for data management of the LeARN project.

Authors' disclosures available online (<http://j-alz.com/manuscript-disclosures/16-0126r2>).

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: <http://dx.doi.org/10.3233/JAD-160126>.

REFERENCES

- [1] De Lepeleire J, Wind AW, Iliffe S, Moniz-Cook ED, Wilcock J, Gonzalez VM, Derksen E, Gianelli MV, Vernooij-Dassen M (2008) The primary care diagnosis of dementia in Europe: An analysis using multidisciplinary, multinational expert groups. *Aging Ment Health* **12**, 568-576.
- [2] Woods RT, Moniz-Cook E, Iliffe S, Campion P, Vernooij-Dassen M, Zanetti O, Franco M (2003) Dementia: Issues in early recognition and intervention in primary care. *J R Soc Med* **96**, 320-324.
- [3] World Health Organization (2012) Dementia: A public health priority, WHO, Geneva, Switzerland.
- [4] Salloway S, Correia S (2009) Alzheimer disease: Time to improve its diagnosis and treatment. *Cleve Clin J Med* **76**, 49-58.
- [5] Galton CJ, Patterson K, Xuereb JH, Hodges JR (2000) Atypical and typical presentations of Alzheimer's disease: A clinical, neuropsychological, neuroimaging and pathological study of 13 cases. *Brain* **123 Pt 3**, 484-498.
- [6] Backman L, Jones S, Berger AK, Laukka EJ, Small BJ (2005) Cognitive impairment in preclinical Alzheimer's disease: A meta-analysis. *Neuropsychology* **19**, 520-531.
- [7] Beck IR, Schmid NS, Berres M, Monsch AU (2014) Establishing robust cognitive dimensions for characterization and differentiation of patients with Alzheimer's disease, mild cognitive impairment, frontotemporal dementia and depression. *Int J Geriatr Psychiatry* **29**, 624-634.
- [8] Ballard C, O'Brien J, Gray A, Cormack F, Ayre G, Rowan E, Thompson P, Bucks R, McKeith I, Walker M, Tovee M (2001) Attention and fluctuating attention in patients with dementia with Lewy bodies and Alzheimer disease. *Arch Neurol* **58**, 977-982.
- [9] Gainotti G, Marra C, Villa G (2001) A double dissociation between accuracy and time of execution on attentional tasks in Alzheimer's disease and multi-infarct dementia. *Brain* **124**, 731-738.
- [10] Tierney MC, Black SE, Szalai JP, Snow WG, Fisher RH, Nadon G, Chui HC (2001) Recognition memory and verbal fluency differentiate probable Alzheimer disease from subcortical ischemic vascular dementia. *Arch Neurol* **58**, 1654-1659.
- [11] American Psychiatric Association (2013) *Diagnostic and statistical manual of mental disorders: DSM-5*, American Psychiatric Association, Washington, DC.
- [12] Handels RL, Aalten P, Wolfs CA, OldeRikkert M, Scheltens P, Visser PJ, Joore MA, Severens JL, Verhey FR (2012) Diagnostic and economic evaluation of new biomarkers for Alzheimer's disease: The research protocol of a prospective cohort study. *BMC Neurol* **12**, 72.
- [13] Folstein MF, Folstein SE, McHugh PR (1975) Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* **12**, 189-198.
- [14] Perneczky R, Wagenpfeil S, Komossa K, Grimmer T, Diehl J, Kurz A (2006) Mapping scores onto stages: Mini-mental state examination and clinical dementia rating. *Am J Geriatr Psychiatry* **14**, 139-144.
- [15] Morris JC (1993) The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology* **43**, 2412-2414.
- [16] Almeida OP, Almeida SA (1999) Short versions of the geriatric depression scale: A study of their validity for the diagnosis of a major depressive episode according to ICD-10 and DSM-IV. *Int J Geriatr Psychiatry* **14**, 858-865.
- [17] Cummings JL (1997) The Neuropsychiatric Inventory: Assessing psychopathology in dementia patients. *Neurology* **48**, S10-S16.
- [18] Gelinas I, Gauthier L, McIntyre M, Gauthier S (1999) Development of a functional measure for persons with Alzheimer's disease: The disability assessment for dementia. *Am J Occup Ther* **53**, 471-481.
- [19] Aalten P, Ramakers I, Biessels G, de Deyn P, Koek HL, OldeRikkert M, Oleksik AM, Richard E, Smits LL, van Swieten JC, Teune LK, van der Lugt A, Barkhof F, Teunissen CE, Rozendaal N, Verhey F, van der Flier WM (2014) The Dutch Parelstoer Institute - Neurodegenerative diseases; methods, design and baseline results. *BMC Neurol* **14**, 1060.
- [20] Brand N, Jolles J (1985) Learning and retrieval rate of words presented auditorily and visually. *J Gen Psychol* **112**, 201-210.
- [21] Lindeboom J, Schmand B, Tulner L, Walstra G, Jonker C (2002) Visual association test to detect early dementia

- of the Alzheimer type. *J Neurol Neurosurg Psychiatry* **73**, 126-133.
- [22] Wechsler D (1997) *Wechsler Memory Scale - 3rd edition: Administration and Scoring Manual*. The Psychological Corporation, San Antonio.
- [23] Lezak M (1995) *Neuropsychological Assessment. 3rd edition.*, Oxford University Press, New York.
- [24] van der Elst W, van Boxtel MP, van Breukelen GJ, Jolles J (2006) The Letter Digit Substitution Test: Normative data for 1,858 healthy participants aged 24-81 from the Maastricht Aging Study (MAAS): Influence of age, education, and sex. *J Clin Exp Neuropsychol* **28**, 998-1009.
- [25] Stroop JR (1935) Studies of interference in serial verbal reactions. *J Exp Psychol* **18**, 643-662.
- [26] Reitan RM (1958) Validity of the Trail Making Test as an indicator of organic brain damage. *Percept Mot Skills* **8**, 271-276.
- [27] Lezak MD, Howieson DB, Loring DW, Hannay HJ, Fischer JS (2004) *Neuropsychological Assessment 4th edition*, Oxford University Press, Oxford.
- [28] Scheltens P, Leys D, Barkhof F, Huglo D, Weinstein HC, Vermersch P, Kuiper M, Steinling M, Wolters EC, Valk J (1992) Atrophy of medial temporal lobes on MRI in probable Alzheimer's disease and normal ageing: Diagnostic value and neuropsychological correlates. *J Neurol Neurosurg Psychiatry* **55**, 967-972.
- [29] Fazekas F, Barkhof F, Wahlund LO, Pantoni L, Erkinjuntti T, Scheltens P, Schmidt R (2002) CT and MRI rating of white matter lesions. *Cerebrovasc Dis* **13**(Suppl 2), 31-36.
- [30] Handels RL, Wolfs CA, Aalten P, Bossuyt PM, Joore MA, Leentjens AF, Severens JL, Verhey FR (2014) Optimizing the use of expert panel reference diagnoses in diagnostic studies of multidimensional syndromes. *BMC Neurol* **14**, 190.
- [31] Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. (2008) Evaluating the added predictive ability of a new marker: From area under the ROC curve to reclassification and beyond. *Stat Med* **27**, 157-172; discussion 207-112.
- [32] Salmon DP, Bondi MW (2009) Neuropsychological assessment of dementia. *Annu Rev Psychol* **60**, 257-282.
- [33] Visser PJ, Verhey F, Knol DL, Scheltens P, Wahlund LO, Freund-Levi Y, Tsolaki M, Minthon L, Wallin AK, Hampel H, Burger K, Pirttila T, Soininen H, Rikkert MO, Verbeek MM, Spuru L, Blennow K (2009) Prevalence and prognostic value of CSF markers of Alzheimer's disease pathology in patients with subjective cognitive impairment or mild cognitive impairment in the DESCRIPA study: A prospective cohort study. *Lancet Neurol* **8**, 619-627.
- [34] Chui H, Zhang Q (1997) Evaluation of dementia: A systematic study of the usefulness of the American Academy of Neurology's practice parameters. *Neurology* **49**, 925-935.
- [35] Meijls AP, Claassen JA, Rikkert MG, Schalk BW, Meulenbroek O, Kessels RP, Melis RJ (2015) How does additional diagnostic testing influence the initial diagnosis in patients with cognitive complaints in a memory clinic setting? *Age Ageing* **44**, 72-77.
- [36] Hentschel F, Kreis M, Damian M, Krumm B, Frolich L (2005) The clinical utility of structural neuroimaging with MRI for diagnosis and differential diagnosis of dementia: A memory clinic study. *Int J Geriatr Psychiatry* **20**, 645-650.
- [37] Bertens LC, van Mourik Y, Rutten FH, Cramer MJ, Lammers JW, Hoes AW, Reitsma JB, Moons KG (2015) Staged decision making was an attractive alternative to a plenary approach in panel diagnosis as reference standard. *J Clin Epidemiol* **68**, 418-425.
- [38] Leening MJ, Vedder MM, Witteman JC, Pencina MJ, Steyerberg EW (2014) Net reclassification improvement: Computation, interpretation, and controversies: A literature review and clinician's guide. *Ann Intern Med* **160**, 122-131.
- [39] Harvey PD (2012) Clinical applications of neuropsychological assessment. *Dialogues Clin Neurosci* **14**, 91-99.
- [40] Noel-Storr AH, McCleery JM, Richard E, Ritchie CW, Flicker L, Cullum SJ, Davis D, Quinn TJ, Hyde C, Rutjes AW, Smailagic N, Marcus S, Black S, Blennow K, Brayne C, Fiorivanti M, Johnson JK, Kopke S, Schneider LS, Simmons A, Mattsson N, Zetterberg H, Bossuyt PM, Wilcock G, McShane R (2014) Reporting standards for studies of diagnostic test accuracy in dementia: The STARDdem Initiative. *Neurology* **83**, 364-373.